

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-305391

(43)Date of publication of application : 22. 11. 1996

(51)Int. Cl. G10L 3/00
G10L 3/00

(21)Application number : 07-106365 (71)Applicant : FUJITSU TEN LTD
(22)Date of filing : 28. 04. 1995 (72)Inventor : TAKAHASHI IKUE
SAKO KAZUYA
FUJIMOTO SHOJI
FUJIMOTO HIROYUKI

(54) VOICE PROCESSING DEVICE

(57)Abstract:

PURPOSE: To simply discriminate selection of kinds of the first (mode) and the second (command) and to operate the device by constituting the device so that it is guided which selecting operation is to be performed depending on sound or display of a guide means.

CONSTITUTION: It is judges whether a switch 7 is pressed for the first time or the second timewhen it is judged that the switch 7 is pressed for the first timea beep sound demanding the first kind selection (mode)for examplea signal sound such as 'pee' is outputted from a loudspeaker 6. At the same timea microcomputer 3 judges that this operation is performed for the first time and the device is in a mode stateand stores the information in a RAM in the microcomputer 3. Nextwhen the switch 7 is operated againthis is detected as the second time operation and the microcomputer 3 judges that the device is in a command stateand stores it in a RAMwhile a beep sound demanding the second kind selection (command)for examplea signal sound such as 'buu' is outputted from a loudspeaker 6. Therebyit can be recognized that the device is in a state in which a command can be inputted.

CLAIMS

[Claim(s)]

[Claim 1] A speech processing unit wherein said operation means is provided with guide mechanism which it shows to said activity which should be chosen according to a predetermined input state by sound or display in a speech processing unit characterized by comprising the following.

An operation start switch.

Inside of at least two kinds of activity which should be chosen from the exterior after detecting operation of this operation start switch Recognize the contents of the sound corresponding to activity of the 1st kind and it operates by activity of the 1st kind based on the contents of the this recognized sound An operation means which recognizes the contents of the sound corresponding to activity of the 2nd kind among said at least two kinds of activity and operates by activity of the 2nd kind based on the contents of the this recognized sound after operating by said 1st kind of activity and detecting operation of said operation start switch.

[Claim 2] The speech processing unit according to claim 1 wherein said guide mechanism is that to which it shows said activity which should be chosen based on detection of operation of said operation start switch by sound or display.

[Claim 3] The speech processing unit according to claim 1 when said guide mechanism has been recognized [the specific contents of a voice] wherein it is what is guided by said activity which should be chosen a sound or display.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to voice recognition equipment and relates to the device which performs the selection instructing for example mode select directions and select command directions of activity with a sound.

[0002]

[Description of the Prior Art] Voice recognition equipment is provided in electronic equipment in recent years and the art which controls apparatus by this speech recognition has been established. Drawing 7 is a flow chart which shows the processing operation of the microcomputer of the

conventional speech processing unit. The processing operation of a microcomputer is started by turning on a speech processing unit. In Step S50 if it judges whether the start switch is pushed and the switch is pushed it will move to Step S51. When the switch is pushed and there is nothing it stands by until it returns at the start time and a switch is pushed.

[0003] In Step S51 the "PITSU" sound which turns into an informing sound is generated and he is urged to give a driver the next directions with a sound. Next it moves to Step S52 and the instruction voice from a driver is incorporated and recognized and it moves to Step S53. And at Step S53 the contents of a voice incorporated and recognized at Step S52 judge the 1st kind (mode) and the 2nd kind (command).

[0004] If judgment at Step S53 is judged to be the 1st kind (mode) it moves to Step S54 and if it is judged as the 2nd kind (command) it will move to Step S55. When the recognized contents of a voice are not what [of the 1st kind (mode) or the 2nd kind (command)] one either it returns and stands by to Step S50. If it is judged as the 1st kind (mode) and moves to Step S54 the signal which directs operation of mode select based on the contents of a voice of mode select will be outputted and processing will be completed. If it is judged as the 2nd kind (command) and moves to Step S55 the signal which directs operation of a select command based on the contents of a voice of a select command will be outputted and processing will be completed.

[0005]

[Problem(s) to be Solved by the Invention] Thus said informing sound is the always same sound and because voice input is only urged he has stopped at the conventional device. Therefore a user may not understand anymore what the next selection operation was when selection of the 1st kind (mode) or selection of the 2nd kind (command) forgets the contents of the last selection operation unintelligibly. In particular if it is a device for mount it cannot concentrate on operation.

[0006] For this reason in this invention selection of the 1st kind (mode) or selection of the 2nd kind (command) distinguishes simply operation is possible and it aims preventing what a deed operating procedure does not understand the next operation display anymore for etc. at providing the speech processing unit which is easy to operate it.

[0007]

[Means for Solving the Problem] In order that this invention may solve such a problem after detecting operation of an operation start switch and this operation start switch The contents of the sound corresponding to activity of the 1st kind are recognized among at least two kinds of

activity which should be chosen from the exteriorAfter operating by activity of the 1st kind based on the contents of the recognized this sound and operating by said 1st kind of activityInside of at least two kinds of said activity after detecting operation of said operation start switchIn a speech processing unit provided with an operation means which recognizes the contents of the sound corresponding to activity of the 2nd kindand operates by activity of the 2nd kind based on the contents of the this recognized soundSaid operation means is provided with guide mechanism which it shows to said activity which should be chosen according to a predetermined input state by sound or display.

[0008]Said guide mechanism is characterized by being that to which it shows said activity which should be chosen based on detection of operation of said operation start switch by sound or display. Said guide mechanism is characterized by being what is guided by said activity which should be chosena soundor displaywhen the specific contents of a voice have been recognized.

[0009]

[Function]When performing selection (mode) of the 1st kindand selection (command) of the 2nd kind by the above means according to this inventionit is shown of what one guide mechanism should just perform selection operation by sounds (a beep soundvoice synthesisetc.) or display.

[0010]

[Example]Hereafterthe example of this invention is described in detail using a drawing. Drawing 1 is a figure showing the composition of one example of this invention. 1 detects an operator's sound with a microphone and inputs the contents of a voice into the voice recognition equipment 2. 2 is voice recognition equipmentit recognizes the audio analog signal inputted from the microphone 1processes it to a digital signaland is outputted to the microcomputer 3. 3 is a microcomputer and processes the digital signal inputted from the voice recognition equipment 2. For examplethe contents of a voice and the contents of a voice of control lead which were memorized as a dictionary are beforehand compared with the microcomputer 3an indication signal is outputted or an indication signal is outputted to each display (LED8HUD9) or the driver 4.

[0011]With directions of the processed signala sound is compounded and it outputs from the loudspeaker 6. 4 is a driver and controls the operation target functions 5 (for exampleoperation of a windowaudio equipmentand a navigation deviceetc.) which operated with the indication signal from the microcomputer 3and were carried in vehicles. 6 is a

loudspeaker changes the audio signal outputted from the microcomputer 3 to a sound and outputs it. 7 is a touch-on switch which is touched to start a speech processing unit and starts this device.

[0012] It is LED 8 and 8 is a display for indication by the light emitting diode which displays the operation situation of a speech processing unit and it switches on the light and blinks with the indication signal from the microcomputer 3 and it displays operation. 9 is HUD (HUD) and by vehicle it is a display projected on a windshield and as the feature the driver can see a display without changing the eye line for operation.

[0013] Next the schematic diagram of drawing 2 explains memory of microcomputer 3 inside. For example, the memory content of ROM (lead on-memory). The mode item from which drawing 2 chooses the activity (mode) of the 1st kind as a horizontal axis is displayed. That is, it is an item of a mounted device and six kinds of mounted devices (an audio, TV navigation, a window, a wiper, a blinker) can be chosen, for example. This is equivalent to the contents of a voice in the mode (an "audio", "TV navigation", a "window", a "wiper", a "blinker"). And the command item which chooses the activity (command) of the 2nd kind is displayed on the vertical axis. That is, it is an operation item of the mounted device selected by the 1st kind. For example, they are ON-OFF of a power supply, adjustment of volume, rise and fall of a window glass, etc. This is equivalent to the contents of a voice in a command (a "power turn", "power OFF", a volume rise, a "rise", etc.).

[0014] These items are compared with the contents of a voice which stored predetermined language in the microcomputer 3 beforehand and it was inputted from the microphone 1 and have been recognized with voice recognition equipment. If it judges that it agreed with the predetermined language which was read from this ROM and which was made to memorize beforehand, the control signal according to the contents of a voice will be outputted to the driver 4 from the microcomputer 3 and the function 5 of an operation target will be operated. A control signal is inputted into LED 8 or HUD 9 and the situation of a control action is displayed.

[0015] When the mode is chosen, the command operation which should recognize the item (for example, window) in the mode in which it was selected in the mode and should be chosen as the next is prepared. At the time of a select command, as only the item (a "rise", "descent", "stop") of the command corresponding to the item in the mode recognized beforehand was extracted, reading processing is simplified.

[0016] Next, the flow chart of drawing 3 explains the 1st example. The processing operation of the microcomputer 3 is started by turning on a speech processing unit. In Step S1, if it judges whether the switch 7 was

pushed on the 1st time or it was pushed on the 2nd time and judges that the 1st switch 7 is pushed it will move to Step S2. If it judges that the switch 7 was pushed on the 2nd time it will move to Step S3.

[0017] When the switch is not pushed it stands by until it returns at the start time and the switch 7 is pushed. In Step S2 the beep sound to which selection (mode) of the 1st kind is urged for example PITSU and a beep is outputted from the loudspeaker 6. It judges that the microcomputer 3 is the 1st operation and is in a mode state with this and RAM (random access memory) in the microcomputer 3 which is not illustrated is made to memorize the information which shows that it is a mode state.

[0018] The sound from a driver is inputted from the microphone 1 after that the sound is recognized by step S4 and it moves to Step S5. Step S5 is a step which judges whether the recognized contents of a voice are the contents of a voice in the mode or they are the contents of a voice in a command and whether they are also the mode or a command. In this case since it is in a mode state the contents of a voice equivalent to a mode item are read and it compares with the contents of a voice recognized this time. Therefore if the recognized contents of a voice have agreed in one of the mode items it will move to Step S6. In differing from the mode the contents of a voice equivalent to a command item etc. return to step S4 and wait for the next spoken command. In Step S6 it will be in the mode state which outputted and asked for the indication signal of the mode operation according to the selected mode and processing is completed.

[0019] Next if the switch 7 is again operated so that a driver may perform operation to a window this is detected as the 2nd operation and it is judged as a command state and this is memorized to RAM and the beep sound to which selection (command) of the 2nd kind is urged for example PUTSU and a beep is outputted from the loudspeaker 6. Thereby the driver differs in the 2nd sound and can recognize that it is in the state of inputting a command.

[0020] Next it moves to step S4 and the contents of a voice are recognized. It is judged whether the contents of a voice recognized at Step S5 are the contents of a voice in a command. That is since it is in a command state at this time the contents of a voice equivalent to a command item are read (in this case a "rise" descent a "stop") and it compares with the contents of a voice recognized this time. Therefore if the contents of a voice agree in one of the command items in predetermined mode (window) it will move to Step S7. When it differs from a command the contents of a voice equivalent to a mode item etc. return to step S4 and wait for the next spoken command.

[0021]At Step S7a signal is outputted to a window actuator in order to make it go up and down an output i.e. a window for the indication signal of command operation according to the selected command item (a "rise" descent stop). this -- a window -- directions -- how -- ***** is carried out. In Step S7 the operation frequency (the 2nd time) of Step S7 which the microcomputer 3 has recognized is reset. The command over the same mode is continued and it may enable it to utter.

[0022]The 2nd example is what is displayed until it is told from an ear to a peep sound and selection operation completes the display by viewing Especially in the case of a car etc. it is the display by HUD (HUD) and as shown in drawing 2 even if the windshield 10 is made to reflect a character or numerals and a driver does not change an eye line displays are the display which enabled it to drive safely by [as I am easy and being able to view].

[0023]Also in the 2nd example the processing operation of the microcomputer 3 is fundamentally the same as the 1st example. However in the example display control is added to processing of Step S2 and Step S3. A control signal is outputted to HUD9 at the same time it will pronounce the sound "PITSU" sound to which mode select is urged at Step S2 if the switch 7 is pushed in Step S1 of the flow chart of drawing 2. The display 12 for which a select command is urged to the windshield 10 which this shows to drawing 4 projects and a control signal is outputted to HUD9 at the same time it pronounces the sound "PUTSU" sound to which a select command is urged at Step S3. It is displayed until the display 13 to which mode select is urged projects on the windshield 10 by this and it starts the next operation.

[0024]Thus as the 2nd example it can distinguish from the difference in the sound by a peep sound by displays simultaneously since the display of HUD is displayed until it completes the next selection selection in the mode or selection of a command understands it clearly next. Although it has shown around by the sound and the display in this example it may show around only especially by display. Drawing 5 is the 3rd example the appearance of the display by two light emitting diodes is shown and the processing operation of the microcomputer 3 is fundamentally [as the 1st example] the same also in the 3rd example. However display control is added to processing of Step S2 and Step S3. Said display is added to for example TV monitor etc.

[0025]The light emitting diode 21 (mode display) of the 1st display and the light emitting diode 22 (command table Shimesu) of the 2nd display are formed into the display 20 which drawing 5 shows the action indication of a speech processing unit The light emitting diode 21 which

displays mode select blinks at the same time it pronounces the "PITSU" sound to which the switch 7 is pushed in Step S1 of the flow chart of drawing 2 it is judged as the 1st time and mode select is urged at Step S2. [0026] And if the contents of a voice of mode select are recognized and mode select is completed the light is put out and the light emitting diode 21 will pronounce the "PITSU" sound to which the switch 7 is pushed next it is judged as the 2nd time and a select command is urged at Step S3 and the light emitting diode 22 which shows the select command will blink it. And if the contents of a voice of the command which the command was directed and was directed are recognized and the selection operation of a command is completed the light emitting diode 22 will be switched off.

[0027] Thus in the 3rd example both the situations where selection operation is performed by blink and putting out lights of the light emitting diode simultaneously with the peep sound by an acoustic sense are also displayed. Although it has shown around by the sound and the display in this example it may show around only especially by display. Next drawing 6 is the 4th example is a display which displays the kind of mode select and displays the 3rd example still in detail. 30 is the display (display) and the kind in mode which is in the display 30 at the mode column of the schematic diagram of drawing 2 for example is displayed (three kinds an audio navigation and a vehicle are displayed in drawing 5).

[0028] For example as shown in drawing 6 31 is a lamp which blinks when an audio is chosen and when 32 chooses a navigation device 33 blinks at the time of the selection which controls rise and fall of the door window of a car and opening and closing of the locking mechanism of a door. Therefore it understands which kind it chose by selection of current mode. At the time of blink of such a lamp if the microcomputer 3 recognizes the contents of a voice of the command corresponding to the mode the microcomputer 3 will make the lamp 31 or the lamp 32 and the lamp 33 turn on the present command will input it and a user (driver) will be told about a device being working.

[0029] Thus the more convenient speech processing unit by an acoustic sense and vision can be provided with combination with the acoustic sense by a peep sound HUD a light emitting diode a mode display lamp etc. Voice synthesis for example choose a command may be used instead of this peep sound. Although it has shown around by the sound and the display in this example it may show around only especially by display.

[0030] Next correspondence in case the item chosen in the 1st example as the 5th example is not yet known (when it chose which kind was chosen or

you have forgotten in the passage of time etc.) is explained. Drawing 7 is a flow chart which shows the processing operation of the microcomputer 3 of the 5th example. Fundamental operation is the same as the 1st example.

[0031]The processing operation of the microcomputer 3 is started by turning on a speech processing unit. In Step S1 it judges whether the switch 7 is pushed and the switch 7 is pushed it will move to Step S2. When the switch 7 is pushed and there is nothing it stands by until it returns at the start time and the switch 7 is pushed. In Step S2 the beep sound to which mode select is urged is emitted. It is judged whether it is the language which has recognized the sound which the user uttered by step S4 after that and this sound defined beforehand at the following step S14 and was stored in the microcomputer 3 for example a help. For example it will urge the 2nd selection at Step S16 if it judges that this is a "help" and it moved to Step S15 it detected whether 1st selection (mode) was performed and mode select is completed it utters by "choose a command" and voice synthesis and processing is completed. If the select command is completed for example it will urge the 1st selection at Step S18 it utters by "choose the mode" and voice synthesis and processing is completed.

[0032]When it judges that it is not a "help" at Step S14 processing equivalent to Step S5 of drawing 1 Step S6 and Step S7 is performed. That is the signal which directs the control action of an instruction content which carried out speech recognition at Step S17 is outputted and processing is completed. Thus in the 5th example he is a device which tells the following operation item by making the language which is called "help" and which was defined beforehand recognize and can understand the contents of selection now further. Even if operation is not clear anymore during operation especially it can run in comfort. In this example although shown around by voice synthesis it may show around with a display.

[0033]Although the selected mode is recognized as operation of the activity of the 1st kind and we decided to prepare command operation in the above the 1st thru/or 5th example The mode which was not limited to this for example was chosen shall be recognized the power supply of the electronic equipment equivalent to this mode shall be turned on and command operation (for example volume rise etc.) other than the power supply ON may be performed as operation of the activity of the 2nd kind.

[0034]Three or more kinds of selection in the mode the 1st select command and the 2nd select command may be performed in addition to two kinds of activity for example. It does not need to be limited to the mode

and activity like a command.

[0035]

[Effect of the Invention] Since he can understand whether what is uttered next is the matter or the 2nd selection items of the 1st selection with an acoustic sense or vision for example according to this invention when operating apparatus with a speech processing unit as explained to details above easy and positive operation can be performed. It can operate in comfort without mind being especially taken by operation in the case of the device for mount.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The figure showing the composition of one example of this invention

[Drawing 2] The schematic diagram showing the memory content of microcomputer 3 inside

[Drawing 3] The flow chart which shows the processing operation of the microcomputer 3

[Drawing 4] The lineblock diagram of HUD (HUD) 9

[Drawing 5] The lineblock diagram of the action indication device which displays a selection operation situation

[Drawing 6] The lineblock diagram of the kind display of mode select

[Drawing 7] The flow chart which shows the processing operation of the microcomputer 3 in the 5th example

[Drawing 8] The flow chart which shows the processing operation of the microcomputer of the conventional speech processing unit

[Description of Notations]

- 1 Microphone
 - 2 Voice recognition equipment
 - 3 Microcomputer
 - 4 Driver
 - 5 Operation target function
 - 6 Loudspeaker
 - 7 Switch
 - 8 LED (display for indication)
 - 9 HUD (HUD)
-

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-305391

(43) 公開日 平成8年(1996)11月22日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 7 1		G 1 0 L 3/00	5 7 1 J
	5 5 1			5 5 1 J

審査請求 未請求 請求項の数 3 O L (全 7 頁)

(21) 出願番号 特願平7-106365

(22) 出願日 平成7年(1995)4月28日

(71) 出願人 000237592

富士通テン株式会社

兵庫県神戸市兵庫区御所通1丁目2番28号

(72) 発明者 高橋 育恵

兵庫県神戸市兵庫区御所通1丁目2番28号

富士通テン株式会社内

(72) 発明者 佐古 和也

兵庫県神戸市兵庫区御所通1丁目2番28号

富士通テン株式会社内

(72) 発明者 藤本 昇治

兵庫県神戸市兵庫区御所通1丁目2番28号

富士通テン株式会社内

最終頁に続く

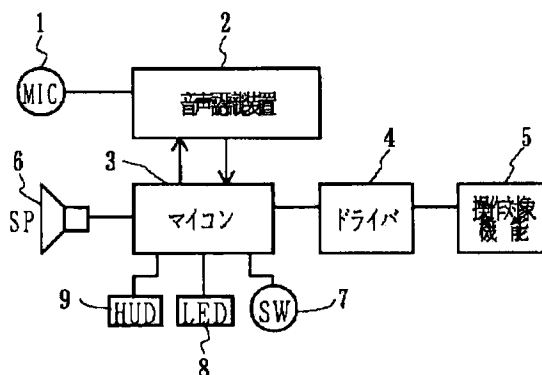
(54) 【発明の名称】 音声処理装置

(57) 【要約】

【目的】 次に選択する動作内容が認識できる音声処理装置を提供する。

【構成】 動作開始スイッチの操作を検出した後、外部から選択されるべき少なくとも2種類の動作内容のうち、第1の種類の動作内容に対応した音声の内容を認識し、認識した音声の内容に基づいて第1の種類の動作内容で動作し、第1の種類の動作内容で動作してから、動作開始スイッチの操作を検出した後、少なくとも2種類の動作内容のうち、第2の種類の動作内容に対応した音声の内容を認識し、認識した音声の内容に基づいて第2の種類の動作内容で動作する動作手段とを備えた音声処理装置に於いて、動作手段は、所定の入力状態に応じて選択されるべき動作内容を音或いは表示にて案内する案内手段を備える。

本発明の一実施例の構成を示す図



【特許請求の範囲】

【請求項 1】 動作開始スイッチと、該動作開始スイッチの操作を検出した後、外部から選択されるべき少なくとも 2 種類の動作内容のうち、第 1 の種類の動作内容に対応した音声の内容を認識し、該認識した音声の内容に基づいて第 1 の種類の動作内容で動作し、前記第 1 の種類の動作内容で動作してから、前記動作開始スイッチの操作を検出した後、前記少なくとも 2 種類の動作内容のうち、第 2 の種類の動作内容に対応した音声の内容を認識し、該認識した音声の内容に基づいて第 2 の種類の動作内容で動作する動作手段とを備えた音声処理装置に於いて、

前記動作手段は、所定の入力状態に応じて選択されるべき前記動作内容を音或いは表示にて案内する案内手段を備えていることを特徴とする音声処理装置。

【請求項 2】 前記案内手段は、前記動作開始スイッチの操作の検出に基づいて選択されるべき前記動作内容を音或いは表示にて案内するものであることを特徴とする請求項 1 記載の音声処理装置。

【請求項 3】 前記案内手段は、特定の音声内容を認識したとき、選択されるべき前記動作内容、音或いは表示で案内するものであることを特徴とする請求項 1 記載の音声処理装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は音声認識装置に係り、動作内容の選択指示、例えばモード選択指示とコマンド選択指示を音声で行う装置に関する。

【0002】

【従来の技術】 近年電子機器には音声認識装置が設けられており、この音声認識によって機器を制御する技術が確立されてきている。図 7 は、従来の音声処理装置のマイコンの処理動作を示すフローチャートである。マイコンの処理動作は、音声処理装置の電源を入れることにより開始される。ステップ S50 では、スタートスイッチが押されているかどうかを判断し、スイッチが押されているとステップ S51 に移る。スイッチが押されていない場合はスタート時点に戻ってスイッチが押されるまで待機する。

【0003】 ステップ S51 では、報知音となる『ピーツ』音を発生させ、運転者に次の指示を音声によって行うように促す。次にステップ S52 に移って運転者からの指示音声を取り込んで認識しステップ S53 に移る。そして、ステップ S53 ではステップ S52 で取り込み認識した音声内容が第 1 の種類（モード）か、第 2 の種類（コマンド）かを判断する。

【0004】 ステップ S53 での判断が第 1 の種類（モード）と判断されれば、ステップ S54 に移り、第 2 の種類（コマンド）と判断するとステップ S55 に移る。認識した音声内容が第 1 の種類（モード）や第 2 の種類

（コマンド）の何方でもない時にはステップ S50 に戻って待機する。第 1 の種類（モード）と判断してステップ S54 に移るとモード選択の音声内容に基づいてモード選択の動作を指示する信号を出力して処理を完了する。又、第 2 の種類（コマンド）と判断してステップ S55 に移るとコマンド選択の音声内容に基づいてコマンド選択の動作を指示する信号を出力して処理を完了する。

【0005】

【発明が解決しようとする課題】 この様に従来の装置では、前記報知音は常時同じ音であり、単に音声入力を促すだけに止まっている。従って、使用者は第 1 の種類（モード）の選択か第 2 の種類（コマンド）の選択かが判り難く、又、前回の選択動作の内容を忘れてしまうことと次の選択動作が何であったか判らなくなってしまうことがある。特に、車載用の装置であれば運転に集中できない。

【0006】 このため本発明においては、第 1 の種類（モード）の選択か第 2 の種類（コマンド）の選択かが簡単に判別して操作ができ、次の操作表示を行い操作手順が判らなくなることを防止する等、操作しやすい音声処理装置を提供することを目的とする。

【0007】

【課題を解決するための手段】 本発明はこの様な問題を解決するために、動作開始スイッチと、該動作開始スイッチの操作を検出した後、外部から選択されるべき少なくとも 2 種類の動作内容のうち、第 1 の種類の動作内容に対応した音声の内容を認識し、該認識した音声の内容に基づいて第 1 の種類の動作内容で動作し、前記第 1 の種類の動作内容で動作してから、前記動作開始スイッチの操作を検出した後、前記少なくとも 2 種類の動作内容のうち、第 2 の種類の動作内容に対応した音声の内容を認識し、該認識した音声の内容に基づいて第 2 の種類の動作内容で動作する動作手段とを備えた音声処理装置に於いて、前記動作手段は、所定の入力状態に応じて選択されるべき前記動作内容を音或いは表示にて案内する案内手段を備えていることを特徴とする。

【0008】 又、前記案内手段は、前記動作開始スイッチの操作の検出に基づいて選択されるべき前記動作内容を音或いは表示にて案内するものであることを特徴とする。又、前記案内手段は、特定の音声内容を認識したとき、選択されるべき前記動作内容、音或いは表示で案内するものであることを特徴とする。

【0009】

【作用】 以上のような手段により本発明によれば、第 1 の種類の選択（モード）と第 2 の種類の選択（コマンド）を行う場合、案内手段が音（ピープ音や音声合成等）或いは表示によって何方の選択動作を行ったか良いかを案内する。

【0010】

【実施例】以下、本発明の実施例について図面を用いて詳細に説明する。図1は、本発明の一実施例の構成を示した図である。1はマイクロホンで操作者の音声を検出し音声内容を音声認識装置2に入力する。2は音声認識装置で、マイクロホン1から入力された音声のアナログ信号を認識しデジタル信号に処理してマイコン3に出力する。3はマイコンで、音声認識装置2から入力されたデジタル信号の処理を行う。又、例えば、予めマイコン3に辞書として記憶した音声内容と制御指示の音声内容を比較して指示信号を出力したり各表示装置（LED8、HUD9）やドライバ4に指示信号を出力する。

【0011】又、処理した信号の指示によって音声を合成してスピーカ6から出力する。4はドライバで、マイコン3からの指示信号により動作して車両に搭載された操作対象機能5（例えばウインド、オーディオ装置、ナビゲーション装置の動作等）を制御する。6はスピーカで、マイコン3から出力される音声信号を音声にかえ出力する。7はタッチONスイッチで、音声処理装置を起動させたい時にタッチして該装置を起動させる。

【0012】8はLEDで、音声処理装置の動作状況を表示する発光ダイオードによる表示器で、マイコン3からの指示信号によって点灯、点滅して動作の表示を行う。9はHUD（ヘッドアップディスプレイ）で、車両などではフロントガラスに映し出される表示装置で、その特徴としては運転者は運転のための視線を変えることなく表示を見ることができる。

【0013】次に図2の概略図によりマイコン3内部の記憶、例えばROM（リード・オンメモリ）の記憶内容について説明する。図2は横軸に第1の種類の動作内容（モード）の選択を行うモード項目が表示されている。即ち、車載装置の項目で、例えば6種類の車載装置（オーディオ、TV、ナビゲーション、ウインド、ワイパー、ウインカ）を選択することができる。これがモードに於ける音声内容（『オーディオ』、『TV』、『ナビゲーション』、『ウインド』、『ワイパ』、『ウインカ』）に相当する。そして、縦の軸には第2の種類の動作内容（コマンド）の選択を行うコマンド項目が表示されている。即ち、第1の種類で選択した車載装置の操作項目である。例えば、電源のON-OFF、音量の調整、ウインドガラスの昇降等である。これがコマンドに於ける音声内容（『電源オン』、『電源オフ』、『音量アップ』、『上昇』等）に相当する。

【0014】これらの項目は所定の言葉を予めマイコン3に記憶させておいて、マイクロホン1から入力されて音声認識装置で認識した音声内容と照合され、このROMから読みだした予め記憶させた所定の言葉と合致したと判断すると音声内容に応じた制御信号がマイコン3からドライバ4に出力されて操作対象の機能5を動作させる。又、LED8やHUD9に制御信号が入力され制御動作の状況の表示を行う。

【0015】尚、モードを選択した時には、そのモードの中で選択されたモードの項目（例えば『ウインド』）を認識して次に選択すべきコマンド動作の準備をしておき、コマンド選択時には、予め認識したモードの項目に対応するコマンドの項目（『上昇』、『下降』、『停止』）だけを抜き出すようにして読み出し処理を簡略化している。

【0016】次に第1の実施例について図3のフローチャートにより説明する。マイコン3の処理動作は、音声処理装置の電源を入れることにより開始される。ステップS1では、スイッチ7が1回目に押されたか、2回目に押されたかを判断し、スイッチ7が1回目押されていると判断するとステップS2に移る。スイッチ7が2回目に押されたと判断するとステップS3に移る。

【0017】又、スイッチが押されていない時は、スタート時点に戻ってスイッチ7が押されるまで待機する。ステップS2では、第1の種類の選択（モード）を促すピープ音、例えば、『ピーッ』と信号音がスピーカ6から出力される。これと共にマイコン3は1回目の操作であり、モード状態にあると判断し、モード状態であることを示す情報をマイコン3内の図示せぬRAM（ランダム・アクセス・メモリ）に記憶させておく。

【0018】その後運転者からの音声をマイクロホン1より入力しステップS4でその音声を認識し、ステップS5に移る。ステップS5は、認識した音声内容がモードにおける音声内容であるか、コマンドにおける音声内容であるか、又モードでもコマンドでもないかの判断を行うステップである。この場合、モード状態にあるのでモード項目に相当する音声内容を読み出し、今回認識した音声内容と比較する。従って認識した音声内容がモード項目の1つに合致しておればステップS6へ移る。コマンド項目に相当する音声内容等、モードと異なる場合にはステップS4に戻って次の音声指示を待つ。ステップS6では、選択されたモードに従ってそのモード動作の指示信号を出力して所望したモード状態となり処理を完了する。

【0019】次に運転者がウインドに対する動作を行うべく、再度スイッチ7を操作すると、これを2回目の操作として検出し、コマンド状態と判断して、RAMにこれを記憶しておくと共に第2の種類の選択（コマンド）を促すピープ音、例えば、『プーッ』と信号音をスピーカ6から出力する。これにより運転者は2回目の音が異なっており、コマンドを入力する状態にあることを認識することができる。

【0020】次にステップS4に移り音声内容を認識する。更にステップS5で認識した音声内容がコマンドにおける音声内容であるか否かを判断する。即ち、この時はコマンド状態にあるのでコマンド項目に相当する音声内容を読み出し（この場合は『上昇』、『下降』、『停止』）、今回認識した音声内容と比較する。従って、音

声内容が所定モード（ウインド）に於けるコマンド項目の1つに合致すればステップS7へ移る。又、モード項目に相当する音声内容等、コマンドと異なる場合はステップS4に戻って次の音声指示を待つ。

【0021】ステップS7では選択されたコマンド項目（『上昇』、『下降』、『停止』）に従いコマンド動作の指示信号を出力、即ちウインドを昇降させるべくウインドアクチュエータに信号を出力する。これによってウインドは指示どおり昇降する。尚、ステップS7に於いて、マイコン3が認識したステップS7の操作回数（2回目）をリセットする。尚、同じモードに対するコマンドを続けて発声する事ができるようにしても良い。

【0022】第2の実施例は、ピープ音で耳から知られると共に目視による表示を選択動作の完了するまで表示するもので、特に自動車の場合などはHUD（ヘッドアップディスプレイ）による表示で、図2に示す如くフロントガラス10に文字又は符号を反射させて運転者が目線を変えなくても表示がたやすく目視できるようにして安全運転が行えるようにした表示装置である。

【0023】第2の実施例においてもマイコン3の処理動作は基本的には第1の実施例と同じである。但し、事例ではステップS2、ステップS3の処理に表示制御を付加している。図2のフローチャートのステップS1においてスイッチ7を押すとステップS2でモード選択を促す音『ピーッ』音を発音すると同時にHUD9へ制御信号を出力する。これにより図4に示すフロントガラス10にコマンド選択を促す表示12が映し出され、ステップS3でコマンド選択を促す音『ブーッ』音を発音すると同時にHUD9へ制御信号を出力する。これによりフロントガラス10にはモード選択を促す表示13が映し出されて次の動作を開始するまでの間表示される。

【0024】この様に第2の実施例としては、ピープ音による音の違いと、表示によって判別することができる。同時にHUDの表示は次の選択を完了するまで表示されるので、次にモードの選択かコマンドの選択かが明確に判る。尚、本例においては音と表示で案内しているが、特に表示のみで案内しても良い。図5は第3の実施例で、発光ダイオード2個による表示装置の外観を示しており、第3の実施例においてもマイコン3の処理動作は第1の実施例と基本的には同じである。但しステップS2、ステップS3の処理に表示制御を付加している。尚、前記表示装置は例えばTVモニター等に付加されている。

【0025】図5は音声処理装置の動作表示を示すディスプレイ20の中に第1表示の発光ダイオード21（モード表示）と第2の表示の発光ダイオード22（コマンド表示）が設けられ、図2のフローチャートのステップS1においてスイッチ7を押して1回目と判断されステップS2でモード選択を促す『ピーッ』音を発音すると同時にモード選択を表示する発光ダイオード21が点滅

する。

【0026】そして、モード選択の音声内容を認識し、モード選択が完了すると発光ダイオード21は消灯して、次にスイッチ7を押して2回目と判断されステップS3でコマンド選択を促す『ブーッ』音を発音すると共にコマンド選択を表示している発光ダイオード22が点滅する。そして、コマンドが指示され指示されたコマンドの音声内容を認識し、コマンドの選択動作が完了すると発光ダイオード22は消灯する。

【0027】この様に、第3の実施例においては、聴覚によるピープ音と同時に発光ダイオードの点滅と消灯によって選択動作が行われている状況も共に表示されるものである。尚、本例においては音と表示で案内しているが、特に表示のみで案内しても良い。次に図6は第4の実施例で、モード選択の種類を表示する表示装置であり第3の実施例の表示を更に詳細に行うものである。30はその表示装置（ディスプレイ）であり、その表示装置30の中に、例えば図2の概略図のモード欄にあるモードの種類が表示される（図5ではオーディオ、ナビゲーション、ピークルの3種類が表示されている）。

【0028】例えば、図6に示すように31はオーディオを選択したときに点滅するランプで、32はナビゲーション装置を選択したとき、33は自動車のドアウインドの昇降やドアのロック機構の開閉を制御する選択時に点滅する。従って、現在モードの選択でどの種類の選択を行ったかが判るようになっている。又、この様なランプの点滅時にマイコン3が、そのモードに対応したコマンドの音声内容を認識すると、マイコン3はランプ31或いはランプ32、ランプ33を点灯させ、現在コマンドが入力して装置が動作中であることを使用者（運転者）に知らせる。

【0029】この様にピープ音による聴覚とHUD、発光ダイオード、モード表示ランプ等との組み合わせによって、聴覚と視覚による、より便利な音声処理装置を提供することができる。尚、このピープ音の代わりに音声合成、例えば『コマンドを選択して下さい』を用いても良い。尚、本例においては音と表示で案内しているが、特に表示のみで案内しても良い。

【0030】次に第5の実施例として、第1の実施例において選択する項目が未だ判らない時（選択を行ったか、どの種類の選択を行ったか、又、時間の経過で忘れた場合など）の対応について説明する。図7は第5の実施例のマイコン3の処理動作を示すフローチャートである。尚、基本的な動作は第1の実施例と同じである。

【0031】マイコン3の処理動作は、音声処理装置の電源を入れることにより開始される。ステップS1では、スイッチ7が押されているかどうかを判断し、スイッチ7が押されているとステップS2に移る。スイッチ7が押されていない場合はスタート時点に戻ってスイッチ7が押されるまで待機する。ステップS2では、モー

ド選択を促すピープ音が発せられる。その後ステップS4で使用者が発した音声を認識し、次のステップS14でこの音声が予め定めてマイコン3に記憶させておいた言葉、例えば『ヘルプ』であるか否かを判断する。これが『ヘルプ』であると判断するとステップS15に移り、第1の選択（モード）を行っていたかどうかを検出して、モード選択が完了しておればステップS16で第2選択を促す、例えば『コマンドを選択して下さい』と音声合成で発声して処理を完了する。又、コマンド選択が完了しておればステップS18で第1選択を促す、例えば『モードを選択して下さい』と音声合成で発声して処理を完了する。

【0032】ステップS14で『ヘルプ』ではないと判断したときには、図1のステップS5、ステップS6、ステップS7に相当する処理を行う。即ち、ステップS17で音声認識した指示内容の制御動作を指示する信号を出力して処理を完了する。この様に第5の実施例においては、『ヘルプ』と言う予め定めた言葉を認識させておくことで次の実施項目を知らせる装置であり、より一層選択内容を理解できるようになっている。特に、運転中に操作が判らなくなっても安心して走行することができる。尚、本例では音声合成で案内しているが表示装置によって案内しても良い。

【0033】以上の第1乃至第5実施例では、第1の種類の動作内容の動作として、選択されたモードを認識し、コマンド動作の準備を行うこととしたが、これに限定されず、例えば選択されたモードを認識し、このモードに相当する電子機器の電源をONするものとし、第2の種類の動作内容の動作として電源ON以外のコマンド動作（例えば音量アップ等）を行うものであっても良い。

【0034】又、2種類の動作内容だけではなく、3種類以上、例えばモードの選択、第1のコマンド選択、第

2のコマンド選択を行うものであっても良い。更にモード、コマンドのような動作内容に限定される必要はない。

【0035】

【発明の効果】以上詳細に説明したように本発明によれば、音声処理装置によって機器の操作を行う場合に、例えば、次に発声するものが第1選択の事項か第2選択事項であるかを聴覚或いは視覚で理解できるために、容易に且つ確実な操作が行える。特に車載用の装置の場合には操作に気をとられることなく安心して運転することができる。

【図面の簡単な説明】

【図1】本発明の一実施例の構成を示す図

【図2】マイコン3内部のメモリ内容を示す概略図

【図3】マイコン3の処理動作を示すフローチャート

【図4】ヘッドアップディスプレイ（HUD）9の構成図

【図5】選択動作状況を表示する動作表示装置の構成図

【図6】モード選択の種類表示装置の構成図

【図7】第5の実施例に於けるマイコン3の処理動作を示すフローチャート

【図8】従来の音声処理装置のマイコンの処理動作を示すフローチャート

【符号の説明】

- 1・・・・・・マイクロホン
- 2・・・・・・音声認識装置
- 3・・・・・・マイコン
- 4・・・・・・ドライバ
- 5・・・・・・操作対象機能
- 6・・・・・・スピーカ
- 7・・・・・・スイッチ
- 8・・・・・・LED（表示器）
- 9・・・・・・HUD（ヘッドアップディスプレイ）

【図2】

マイコン3内部のメモリ内容を示す概略図

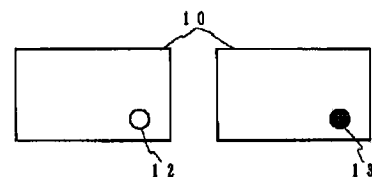
		モ ー ド					
		オーディオ	T	V	ナビゲーション	ウィンド	ワイパー
コ	電源ON	○	○	○			
	電源OFF	○	○	○			
マ	音量+	○	○	○			
	音量-	○	○	○			
ン	上昇		①	③	○	⑤	
	下降		②	④	○	⑥	
ド	停止		○	⑤	○	○	
	右			⑥			○
	左			⑦			○

- ① ch選択上へ
- ② ch選択下へ
- ③ 音量上へ
- ④ 音量下へ
- ⑤ 音量停止

- ⑥ 画面右へ
- ⑦ 画面左へ
- ⑧ ワイパー速度上昇
- ⑨ ワイパー速度下降

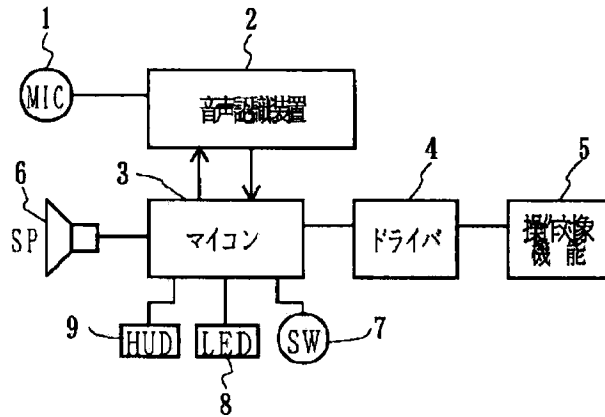
【図4】

ヘッドアップディスプレイ（HUD）9の構成図



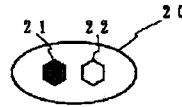
【図1】

本発明の一実施例の構成を示す図



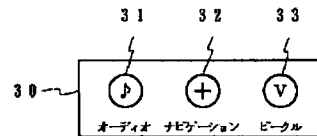
【図5】

選択動作状況を表示する動作表示装置の構成図



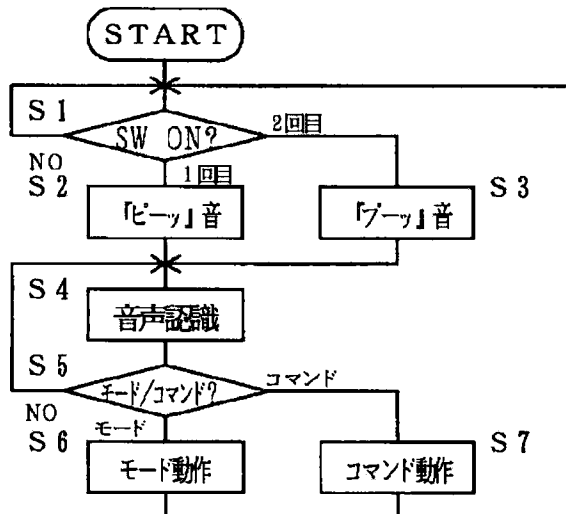
【図6】

選択の種類表示装置の構成図



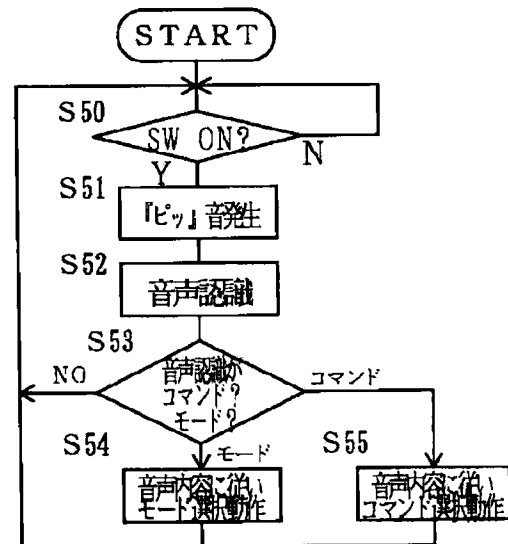
【図3】

マイコン3の処理動作を示すフローチャート



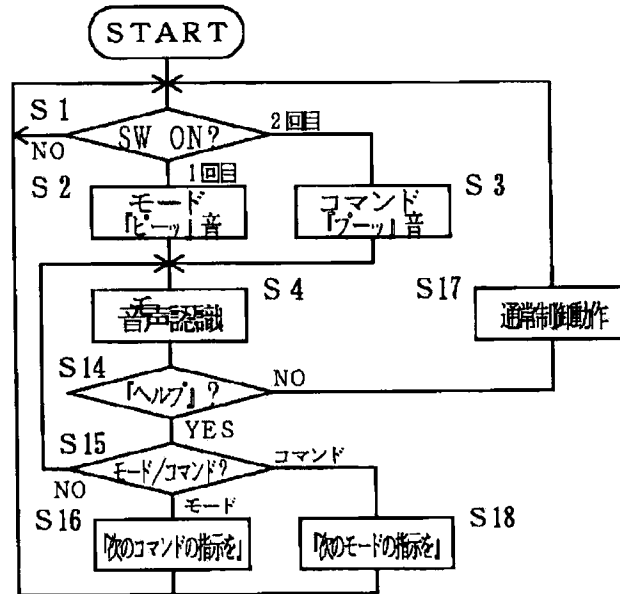
【図8】

従来の音声処理装置のマイコンの処理動作を示すフローチャート



【図7】

第5実施例におけるマイコン3の
処理動作を示すフローチャート



フロントページの続き

(72)発明者 藤本 博之
兵庫県神戸市兵庫区御所通1丁目2番28号
富士通テン株式会社内